AD 616780

MEMORANDUM
RM-4593-PR
JUNE 1965

DYNAMIC PROGRAMMING, SYSTEM IDENTIFICATION, AND SUBOPTIMIZATION

Richard Bellman

17-P
COPY OF
HARD COPY $. 1.00
MICROFICHE $. 0.50

DDC
JUN 29 1965
RECEIVED
DDC-IRA E

PREPARED FOR:

UNITED STATES AIR FORCE PROJECT RAND

The RAND Corporation
SANTA MONICA • CALIFORNIA

ARCHIVE COPY

MEMORANDUM
RM-4593-PR
JUNE 1965

DYNAMIC PROGRAMMING, SYSTEM IDENTIFICATION, AND SUBOPTIMIZATION

Richard Bellman

This research is sponsored by the United States Air Force under Project RAND—Contract No. AF 49(638)-700—monitored by the Directorate of Operational Requirements and Development Plans, Deputy Chief of Staff, Research and Development, Hq USAF. Views or conclusions contained in this Memorandum should not be interpreted as representing the official opinion or policy of the United States Air Force.

DDC AVAILABILITY NOTICE
Qualified requesters may obtain copies of this report from the Defense Documentation Center (DDC).

Approved for OTS release

The RAND Corporation

PREFACE

In this Memorandum the author employs the mathematical technique of dynamic programming to obtain a best-fit approximation to a function that is defined over some given interval. He then describes how this method offers an approach to the handling of a certain type of pattern-recognition problem and to the approximation of optimal control policies.

SUMMARY

The problem we start with appears to be quite specialized. Given a function $u(t)$ defined over the interval $[0,a]$, we wish to find a polygonal approximation which is a best fit in a mean-square sense. The analytic problem for N is that of minimizing the function

$$R_N = \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} (u(t) - a_i - b_i t)^2 dt$$

over the quantities a_i, b_i, and t_i. Here $t_0 = 0$, $t_N = a$.

This can be treated in a number of direct fashions, using search and gradient techniques. We wish, however, to employ dynamic programming, which appears to be superior even in this case, and then gradually to enlarge the scope of the problem until it covers a question in the identification of systems and a version of the general problem of considering suboptimal policies in control processes.

CONTENTS

PREFACE iii

SUMMARY v

Section

1. INTRODUCTION 1

2. ADAPTIVE CURVE FITTING 3

3. DISCUSSION 5

4. IDENTIFICATION OF SYSTEMS 7

5. SUBOPTIMIZATION 9

6. REDUCTION OF DIMENSIONALITY 11

REFERENCES 13

DYNAMIC PROGRAMMING, SYSTEM IDENTIFICATION, AND SUBOPTIMIZATION

1. INTRODUCTION

The problem we start with appears to be quite specialized. Given a function $u(t)$ defined over the interval $[0,a]$, we wish to find a polygonal approximation which is a best fit in a mean-square sense. (See Fig. 1.) The analytic problem for N is that of minimizing the function

$$R_N = \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} (u(t) - a_i - b_i t)^2 dt \tag{1.1}$$

over the quantities a_i, b_i, and t_i. Here $t_0 = 0$, $t_N = a$.

This can be treated in a number of direct fashions, using search and gradient techniques. We wish, however, to employ dynamic programming, which appears to be superior even in this case, and then gradually to enlarge the scope of the problem until it covers a question in the identification of systems and a version of the general problem of considering suboptimal policies in control processes. Results related to what follows have been presented in [1,2,3].

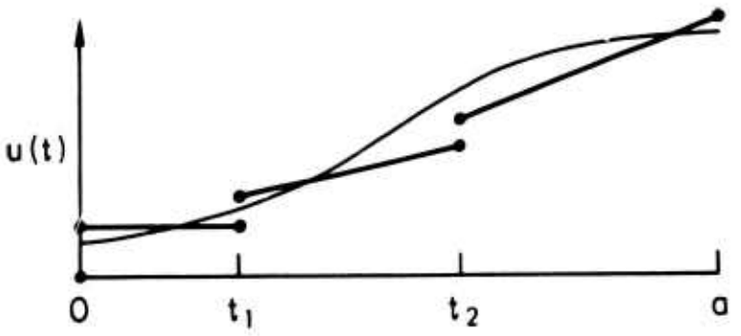

Fig. 1

2. ADAPTIVE CURVE FITTING

The foregoing problem can be considered to fall within the new area of sequential computation. In place of choosing the t_i in advance, we allow the structure of the function $u(t)$ to determine their positions. Similar techniques can be applied in connection with the numerical integration of ordinary and partial differential equations. Write

$$\min_{(a_i,b_i,t_i)} R_N = f_N(a), \tag{2.1}$$

defined for $N = 0,1,2,\ldots,$ and $a \geq 0$. Introduce the function of two variables,

$$\Delta(s_1,s_2) = \min_{a,b} \int_{s_1}^{s_2} (u(t) - a - bt)^2 dt, \tag{2.2}$$

for $0 \leq s_1 \leq s_2 < \infty$. That this happens in this case to be explicitly calculable is of no particular significance at the moment. In general, this function will be obtained via numerical methods.

Then

$$f_0(a) = \Delta(0,a), \tag{2.3}$$

and the principle of optimality yields the recurrence relation

$$(2.4) \qquad f_N(a) = \min_{0 \le t_N \le a} [\Delta(t_N, a) + f_{N-1}(t_N)],$$

for $N \geq 1$.

This leads to a quite simple and efficient computational algorithm.

3. DISCUSSION

Perhaps the first point to note in connection with what has been given above is that the computational feasibility of the algorithm inherent in (2.4) is not strongly dependent upon the mean-square norm in (2.2). We could just as easily use

$$\Delta(s_1,s_2) = \min_{a,b} \max_{s_1 \le t \le s_2} |u(t) - a - bt|, \tag{3.1}$$

or allow approximation by polynomials of higher degree. This brings us into contact with the theory of spline approximations, but we shall not pursue that here; see [4] for an extensive set of references.

As soon as we start pursuing the idea of approximating to $u(t)$ over the interval $[s_1,s_2]$ by a function of simple analytic form, we enter the domain of differential approximation [5]. We recognize that a polynomial of degree M satisfies the differential equation

$$\frac{d^{(M+1)}v}{dt^{M+1}} = 0, \tag{3.2}$$

that the exponential polynomial $\sum_{k=1}^{M} a_k e^{\lambda_k t}$ satisfies the differential equation

$$\frac{d^{(M)}v}{dt^{(M)}} + b_1 \frac{d^{(M-1)}v}{dt^{(M-1)}} + \cdots + b_M v = 0, \tag{3.3}$$

and that $\sum_{k=1}^{M} a_k \cos(\lambda_k + \phi_k)$ satisfies a similar equation of degree 2M. It follows that a substantial extension of straight-line approximation is the following. Determine the parameters a_i and initial conditions c_i so that

$$\|u - v\| \tag{3.4}$$

is minimized, where u is given and v is determined by the ordinary differential equation

$$\frac{d^{(M)}v}{dt^{(M)}} = \left(t, v, \ldots, \frac{d^{(M-1)}v}{dt^{(M-1)}}, a_1, \ldots, a_M\right), \tag{3.5}$$

$v^{(i)}(0) = c_i$, $i = 0,1,\ldots,M-1$. Here, we can use a mean-square norm, or some other convenient norm.

Problems of this nature can be attacked by means of quasilinearization and other techniques [5].

4. IDENTIFICATION OF SYSTEMS

The foregoing remarks and techniques allow us to approach an interesting problem in the identification of systems. Suppose that we know that a function $u(t)$ is generated in the following manner. In the interval $t_i \leq t \leq t_{i+1}$, $t_0 \leq t_1 \leq \cdots \leq t_{n+1}$, $t_0 = 0$, $t_{N+1} = a_0$, it satisfies the equation

$$\frac{d^M v}{dt^{M+1}} = g\left(t, v, \ldots, \frac{d^{(M-1)} v}{dt^{(M-1)}}, a_i\right), \tag{4.1}$$

$$v^{(j)}(t_i) = c_{ij}, \quad j = 0,1,\ldots,M-1.$$

Given the values of $u(t)$ in $[0,a]$, we wish to determine the vector parameters a_i, the parameters c_{ij}, and the switching points t_i, and occasionally N itself. This is a particular type of pattern recognition problem.

We begin by introducing the function

$$\Delta(s_1,s_2) = \min_{a,c_j} \int_{s_1}^{s_2} (u - v)^2 dt, \tag{4.2}$$

where $v(t)$ satisfies (4.1), $0 \leq s_1 \leq s_2 \leq a$. Our assumption is that we can compute this function of two variables. This will in general, however, be a nontrivial task. If then we introduce the function

$$(4.3) \qquad f_N(a) = \min_{\{a_i, c_{ij}\}} \int_0^a (u - v)^2 dt,$$

$a \geq 0$, allowing N switch points, or transition points, we obtain exactly the same recurrence relation as in (2.4). If $u(t)$ is actually determined by (4.1), we will have $f_N(a_0) = 0$ for the correct choice of t_N.

5. SUBOPTIMIZATION

For analytic, economic, and engineering convenience, it is often useful to consider the approximation of optimal control policies by simple, feasible control policies.

Thus, for example, in the minimization of

$$(5.1) \qquad J(u) = \int_0^T g(u,u')dt, \quad u(0) = c,$$

we may wish to consider as admissible functions only those for which

$$(5.2) \qquad u'(t) = b_i, \quad s_i \leq t \leq s_{i+1},$$

with $s_0 = 0$, $s_{N+1} = T$, where the b_i and s_i are to be chosen.

Let us define

$$(5.3) \qquad f_N(T,c) = \min J(u),$$

where the minimum is now over the class of suboptimal policies defined above. Then, as before, the principle of optimality yields the relation

$$(5.4) \qquad f_N(T,c) = \min_{b_0,s_1} \left[\int_0^{s_1} g(u(b_0,t),b_0)dt + f_{N-1}(T - s_1, u(b_0,s_1)) \right],$$

for $N \geq 1$, with

$$f_0(T,c) = \min_{b_0} \int_0^T g(u(b_0,t),b_0)dt.$$

Here $u(b_0,t)$ denotes the function over the relevant t-interval determined by the nature of the suboptimal policy and the initial state c. In this case, $u(b_0,t) = c + b_0 t$.

6. REDUCTION OF DIMENSIONALITY

One of the purposes of using suboptimal policies is to bypass some of the analytic and computational difficulties of the original optimization problem. This is particularly the case when we have a control process involving either a high-dimensional state vector, or an infinite-dimensional vector.

In this situation, we can often replace the actual state vector at time t by a record of the control policies used, and thus obtain a more manageable computational algorithm. Furthermore, we can use new types of approximation methods. For a detailed discussion of this technique, see [6].

REFERENCES

1. Bellman, R., "On the Approximation of Curves by Line Segments Using Dynamic Programming," Comm. Assoc. Comput. Machinery, Vol. 4, 1961, p. 284.

2. Bellman, R., B. Gluss, and R. Roth, "On the Identification of Systems and the Unscrambling of Data: Some Problems Suggested by Neurophysiology," Proc. Nat. Acad. Sci. USA, Vol. 52, 1964, pp. 1239–1249.

3. ———, Segmental Differential Approximation and the "Black Box" Problem, The RAND Corporation, RM-4269-PR, October 1964; to appear in J. Math. Anal. and Appl.

4. Ahlberg, J. H., E. N. Nilson, and J. L. Walsh, "Best Approximation and Convergence Properties of Higher-order Spline Approximations," J. Math. Mech., Vol. 14, 1965, pp. 231–244.

5. Bellman, R., and R. Kalaba, Quasilinearization and Boundary-value Problems, American Elsevier Publishing Company, Inc., New York, to appear in 1965.

6. Bellman, R., Dynamic Programming, Generalized States, and Switching Systems, The RAND Corporation, RM-4474-PR, February 1965; to appear in J. Math. Anal. and Appl.